

H. grisea CBH1.1

Figure 1: Total Genomic Sequence (1638 nucleotides):

1	ATGCGTACCG	CCAAGTTCGC	CACCCTCGCC	GCCCTTGTGG	CCTCGGCCGC	50
51	CGCCCAGCAG	GCGTGACGTC	TCACCACCGA	GAGGCACCCT	TCCCTCTCTT	100
101	GGAAGAAGTG	CACCGCCGGC	GGCCAGTGCC	AGACCGTCCA	GGCTTCCATC	150
151	ACTCTCGACT	CCAAGTGGCG	CTGGACTCAC	CAGGTGTCTG	GCTCCACCAA	200
201	CTGCTACACG	GGCAACAAGT	GGGATACTAG	CATCTGCACT	GATGCCAAGT	250
251	CGTGCGCTCA	GAAGTGTGTC	GTCGATGGTG	CCGACTACAC	CAGCACCTAT	300
301	GGCATCACCA	CCAACGGTGA	TTCCCTGAGC	CTCAAGTTCG	TCACCAAGGG	350
351	CCAGCACTCG	ACCAACGTCG	GCTCGCGTAC	CTACCTGATG	GACGGCGAGG	400
401	ACAAGTATCA	GAGTACGTTT	TATCTTCAGC	CTTCTCGCGC	CTTGAATCCT	450
451	GGCTAACGTT	TACACTTCAC	AGCCTTCGAG	CTCCTCGGCA	ACGAGTTCAC	500
501	CTTCGATGTC	GATGTCTCCA	ACATCGGCTG	CGGTCTCAAC	GGCGCCCTGT	550
551	ACTTCGTCTC	CATGGACGCC	GATGGTGGTC	TCAGCCGCTA	TCCTGGCAAC	600
601	AAGGCTGGTG	CCAAGTACGG	TACCGGCTAC	TGCGATGCTC	AGTGCCCCCG	650
651	TGACATCAAG	TTCATCAACG	GCGAGGCCAA	CATTGAGGGC	TGGACCGGCT	700
701	CCACCAACGA	CCCCAACGCC	GGCGCGGGCC	GCTATGGTAC	CTGCTGCTCT	750
751	GAGATGAGTA	TCTGGGAAGC	CAACAACATG	GCTACTGCCT	TCACTCCTCA	800
801	CCCTTGACCC	ATCATTGGCC	AGAGCCGCTG	CGAGGGCGAC	TCGTGCGGTG	850
851	GCACCTACAG	CAACGAGCGC	TACGCCGGCG	TCTGCGACCC	CGATGGCTGC	900
901	GACTTCAACT	CGTACCGCCA	GGGCAACAAG	ACCTTCTACG	GCAAGGGCAT	950
951	GACCGTCGAC	ACCACCAAGA	AGATCACTGT	CGTCACCCAG	TTCCTCAAGG	1000
1001	ATGCCAACGG	CGATCTCGGC	GAGATCAAGC	GCTTCTACGT	CCAGGATGGC	1050
1051	AAGATCATCC	CCAAGTCCGA	GTCCACCATC	CCCGGCGTCG	AGGGCAATTC	1100
1101	CATCACCAG	GACTGGTGGC	ACCGCCAGAA	GGTTGCCTTT	GGCGACATTG	1150
1151	ACGACTTCAA	CCGCAAGGGC	GGCATGAAGC	AGATGGGCAA	GGCCCTCGCC	1200
1201	GGCCCCATGG	TCCTGGTCAT	GTCCATCTGG	GATGACCACG	CCTCCAACAT	1250
1251	GCTCTGGCTC	GACTCGACCT	TCCCTGTCTG	TGCCGCTGGC	AAGCCCGGCG	1300
1301	CCGAGCGCGG	TGCCTGCCCC	ACCACCTCGG	GTGTCCCTGC	TGAGGTTGAG	1350
1351	GCCGAGGCC	CCAACAGCAA	CGTCGTCTTC	TCCAACATCC	GCTTCGGCCC	1400
1401	CATCGGCTCG	ACCGTTGCTG	GTCTCCCCGG	CGCGGGCAAC	GGCGGCAACA	1450
1451	ACGGCGGCAA	CCCCCGGCC	CCCACCACCA	CCACCTCCTC	GGCTCCGGCC	1500
1501	ACCACCACCA	CCGCCAGCGC	TGGCCCCAAG	GCTGGCCGCT	GGCAGCAGTG	1550
1551	CGGCGGCATC	GGCTTCACTG	GCCCCGACCA	GTGCGAGGAG	CCCTACACTT	1600
1601	GCACCAAGCT	CAACGACTGG	TACTCTCAGT	GCCTGTAA		1638

Figure 2: Putative intron sequence deleted (GTACGTT...CAG = 413-472) Gives the cDNA sequence (1578 nucleotides):

1	ATGCGTACCG	CCAAGTTCGC	CACCCTCGCC	GCCCTTGTTG	CCTCGGCCGC	50
51	CGCCCAGCAG	GCGTGCAATC	TCACCACCGA	GAGGCACCCT	TCCCTCTCTT	100
101	GGAAGAAGTG	CACCGCCGGC	GGCCAGTGCC	AGACCGTCCA	GGCTTCCATC	150
151	ACTCTCGACT	CCAACGGGCG	CTGGACTCAC	CAGGTGTCTG	GCTCCACCAA	200
201	CTGCTACACG	GGCAACAAGT	GGGATACTAG	CATCTGCACT	GATGCCAAGT	250
251	CGTGCGCTCA	GAACGTGCTG	GTCGATGGTG	CCGACTACAC	CAGCACCTAT	300
301	GGCATCACCA	CCAACGGTGA	TTCCCTGAGC	CTCAAGTTCG	TCACCAAGGG	350
351	CCAGCACTCG	ACCAACGTCG	GCTCGCGTAC	CTACCTGATG	GACGGCGAGG	400
401	ACAAGTATCA	GACCTTCGAG	CTCCTCGGCA	ACGAGTTCAC	CTTCGATGTC	450
451	GATGTCTCCA	ACATCGGCTG	CGGTCTCAAC	GGCGCCCTGT	ACTTCGTCTC	500
501	CATGGACGCC	GATGGTGGTC	TCAGCCGCTA	TCCTGGCAAC	AAGGCTGGTG	550
551	CCAAGTACGG	TACCGGCTAC	TGCGATGCTC	AGTGCCCCCG	TGACATCAAG	600
601	TTCATCAACG	GCGAGGCCAA	CATTGAGGGC	TGGACCGGCT	CCACCAACGA	650
651	CCCCAACGCC	GGCGCGGGCC	GCTATGGTAC	CTGCTGCTCT	GAGATGGATA	700
701	TCTGGGAAGC	CAACAACATG	GCTACTGCCT	TACTCCTCA	CCCTTGACAC	750
751	ATCATTTGCC	AGAGCCGCTG	CGAGGGCGAC	TCGTGCGGTG	GCACCTACAG	800
801	CAACGAGCGC	TACGCCGGCG	TCTGCGACCC	CGATGGCTGC	GACTTCAACT	850
851	CGTACCGCCA	GGCAACAAG	ACCTTCTACG	GCAAGGGCAT	GACCGTCGAC	900
901	ACCACCAAGA	AGATCACTGT	CGTCACCCAG	TTCTCAAGG	ATGCCAACGG	950
951	CGATCTCGGC	GAGATCAAGC	GCTTCTACGT	CCAGGATGGC	AAGATCATCC	1000
1001	CCAACTCCGA	GTCCACCATC	CCCGGCGTCG	AGGGCAATTC	CATCACCAG	1050
1051	GACTGGTGCG	ACCGCCAGAA	GGTTGCCTTT	GGCGACATTG	ACGACTTCAA	1100
1101	CCGCAAGGGC	GGCATGAAGC	AGATGGGCAA	GGCCCTCGCC	GGCCCCATGG	1150
1151	TCCTGGTCAT	GTCCATCTGG	GATGACCACG	CCTCCAACAT	GCTCTGGCTC	1200
1201	GACTCGACCT	TCCCTGTCTG	TGCCGCTGGC	AAGCCCGGCG	CCGAGCGCGG	1250
1251	TGCCTGCCCC	ACCACCTCGG	GTGTCCCTGC	TGAGGTTGAG	GCCGAGGCCC	1300
1301	CCAACAGCAA	CGTCGTCTTC	TCCAACATCC	GCTTCGGCCC	CATCGGCTCG	1350
1351	ACCGTTGCTG	GTCTCCCCGG	CGCGGGCAAC	GGCGGCAACA	ACGGCGGCAA	1400
1401	CCCCCGCCCC	CCACACCA	CCACCTCCTC	GGCTCCGGCC	ACCACCA	1450
1451	CCGCCAGCGC	TGGCCCCAAG	GCTGGCCGCT	GGCAGCAGTG	CGGCGGCATC	1500
1501	GGCTTCACTG	GCCCCACCCA	GTGCGAGGAG	CCCTACACTT	GCACCAAGCT	1550
1551	CAACGACTGG	TACTCTCAGT	GCCTGTAA			1578

Figure 3: Translation of the cDNA sequence gives the *H. grisea var thermoidea* CBH1 precursor (i.e.: with signal sequence) protein sequence (525 amino acids):

1	<u>MRTAKFATLA</u>	<u>ALVASAAAQQ</u>	ACSLTTERHP	SLSWKKCTAG	GQCQTVQASI	50
51	TLDSNWRWTH	QVSGSTNCYT	GNKWDTSICT	DAKSCAQNCC	VDGADYTSTY	100
101	GITTNGDSL	LKFVTKQHS	TNVGSRTYLM	DGEDKYQTFE	LLGNEFTFDV	150
151	DVSNIGCGLN	GALYFVSMDA	DGGLSRYPGN	KAGAKYGTGY	CDAQCPRIK	200
201	FINGEANIEG	WTGSTNDPNA	GAGRYGTCCS	EMDIWEANNM	ATAFTPHPCT	250
251	IIGQSRCEGD	SCGGTYSNER	YAGVCDPDGC	DFNSYRQGNK	TFYKGGMTVD	300
301	TTKKITVVVQ	FLKDANGDLG	EIKRFYVQDG	KIIPNSESTI	PGVEGNSITQ	350
351	DWCDRQKQAF	GDIDDFNRKG	GMKQMGKALA	GPMVLVMSIW	DDHASNMLWL	400
401	DSTFPVDAAG	KPGAERGACP	TTSGVPAEVE	AEAPNSNVVF	SNIRFGPIGS	450
451	TVAGLPGAGN	GGNNGGNPPP	PTTTTSSAPA	TTTTASAGPK	AGRWQC CGI	500
501	GFTGPTQCEE	PYTCTKLNDW	YSQCL			525

Figure 4: Mature (i.e.: expressed protein with the putative signal sequence removed) protein sequence (507 amino acids):

1	QQACSLTTER	HPSLSWKKCT	AGGQCQTVQA	SITLDSNWRW	THQVSGSTNC	50
51	YTGKWDTSI	CTDAKSCAQN	CCVDGADYTS	TYGITTNNGDS	LSLKFVTKGQ	100
101	HSTNVGSRTY	LMDGEDKYQT	FELLGNEFTF	DVDVSNIGCG	LNGALYFVSM	150
151	DADGGLSRYP	GNKAGAKYGT	GYCDAQCPRD	IKFINGEANI	EGWTGSTNDP	200
201	NAGAGRYGTC	CSEMDIWEAN	NMATAFTPHP	CTIIGQSRCE	GDSCGGTYSN	250
251	ERYAGVCDPD	GCDFNSYRQG	NKTFYKGGMT	VDTTKKITVV	TQFLKDANGD	300
301	LGEIKRFYVQ	DGKIIPNSES	TIPGVEGNSI	TQDWCDRQKV	AFGDIDDFNR	350
351	KGGMKQMGKA	LAGPMVLVMS	IWDDHASNML	WLDSTFPVDA	AGKPGAERGA	400
401	CPTTSGVPAE	VEAEAPNSNV	VFSNIRFGPI	GSTVAGLPGA	GNGGNNGGNP	450
451	PPPTTTTSSA	PATTTTASAG	PKAGRWQCG	GIGFTGPTQC	EEPYTCTKLN	500
501	DWYSQCL					507

[illegible]

Figure 5. Sequence alignment of two public sequences and variant *H. grisea* CBH1.1

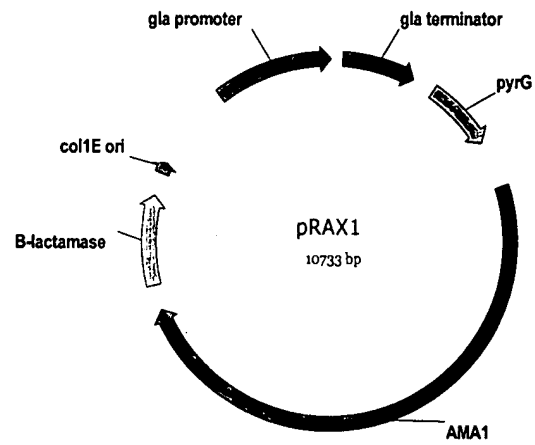


Figure 6: pRAX1

Figure 7: Destination vector pRAXdes2 for expression in *A. niger*

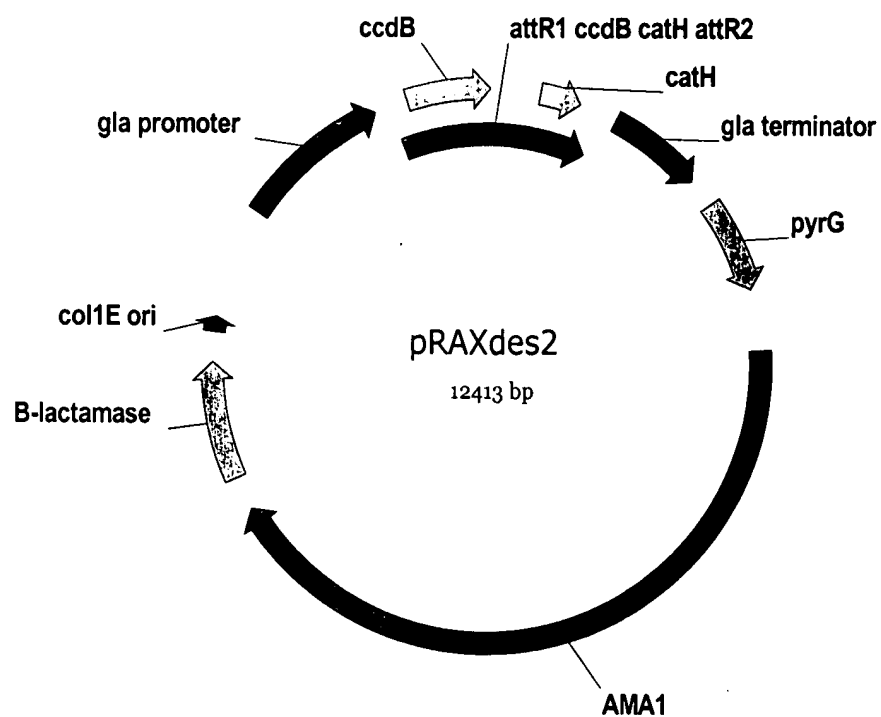


Figure 8: Replicative expression pRAXdesCBH1 vector of CBH1 genes under the control of the glucoamylase promotor.

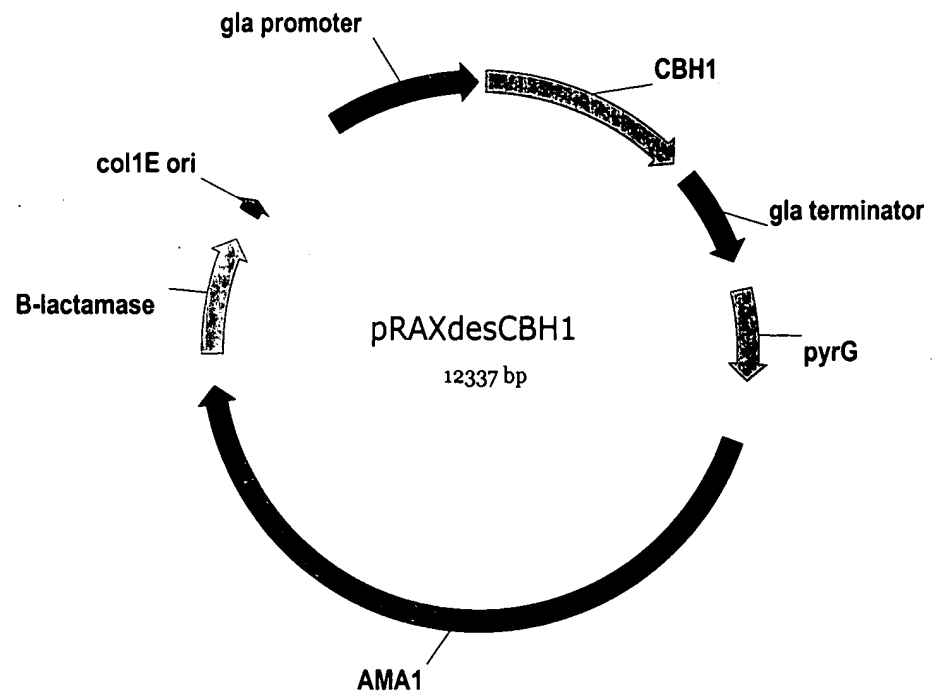


Figure 9: CBHI Conversion Assay:
38C, 700 rpm agitation, 1 day, 12.66% PCS, 15.5 mg enzyme/g cellulose

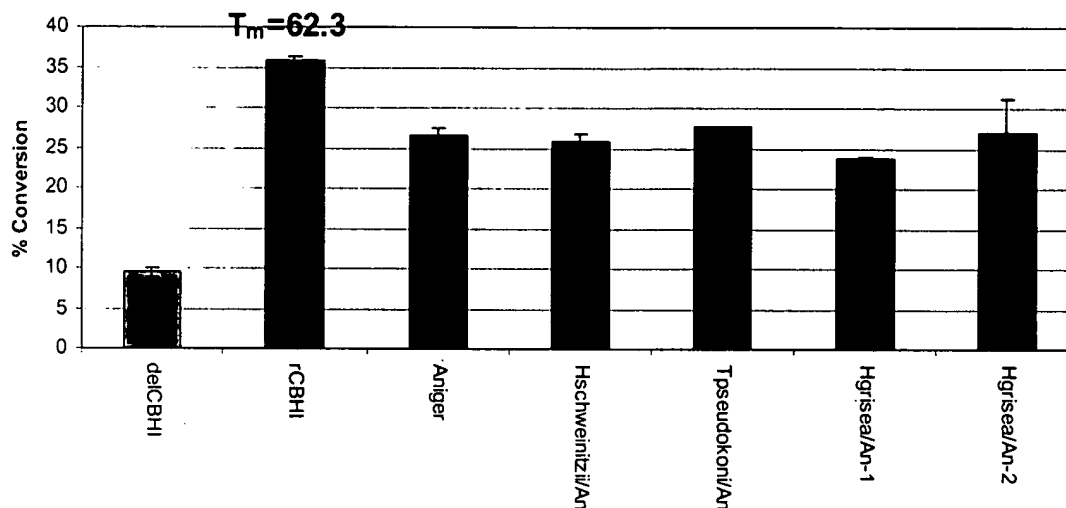


Figure 10: CBHI Conversion Assay:
65C, 700 rpm agitation, 1 day, 12.66% PCS, 15.5 mg enzyme/g cellulose

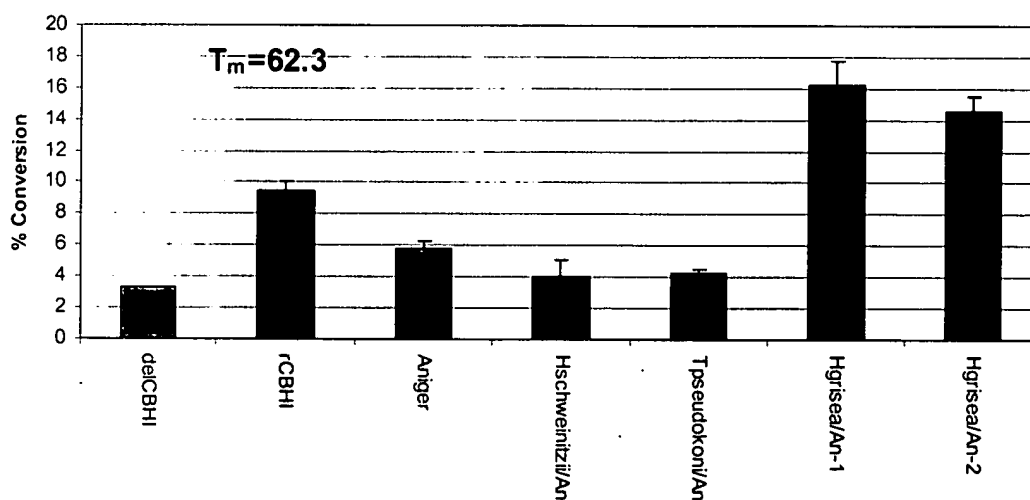


Figure 1 1: CBHI Conversion Assay:
38C, 700 rpm agitation, 1 day, 12.66% PCS, 15.5 mg enzyme/g cellulose

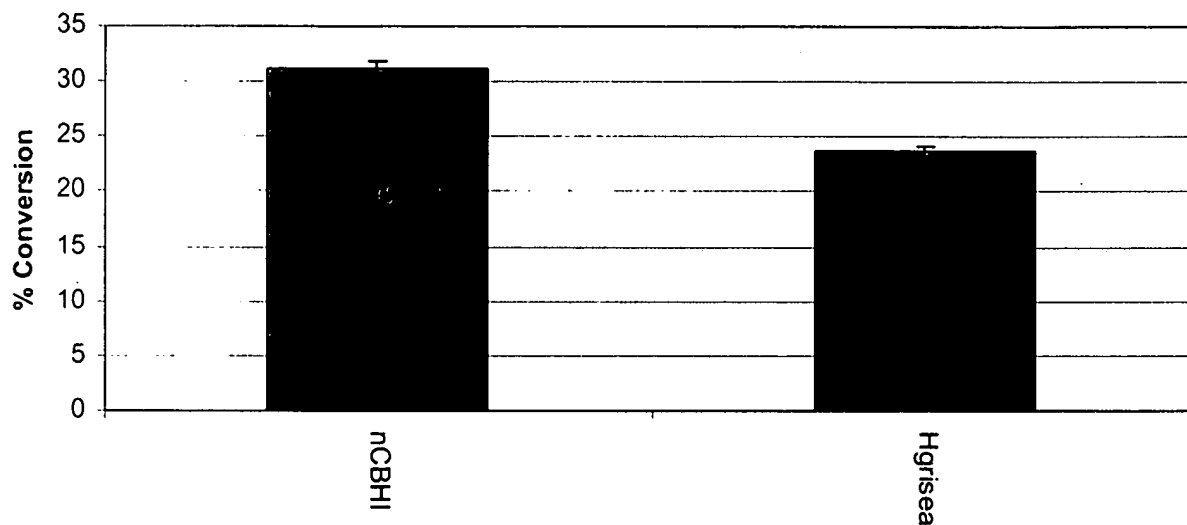


Figure 12: CBHI Conversion Assay:
65C, 700 rpm agitation, 1 day, 12.66% PCS, 15.5 mg enzyme/g cellulose

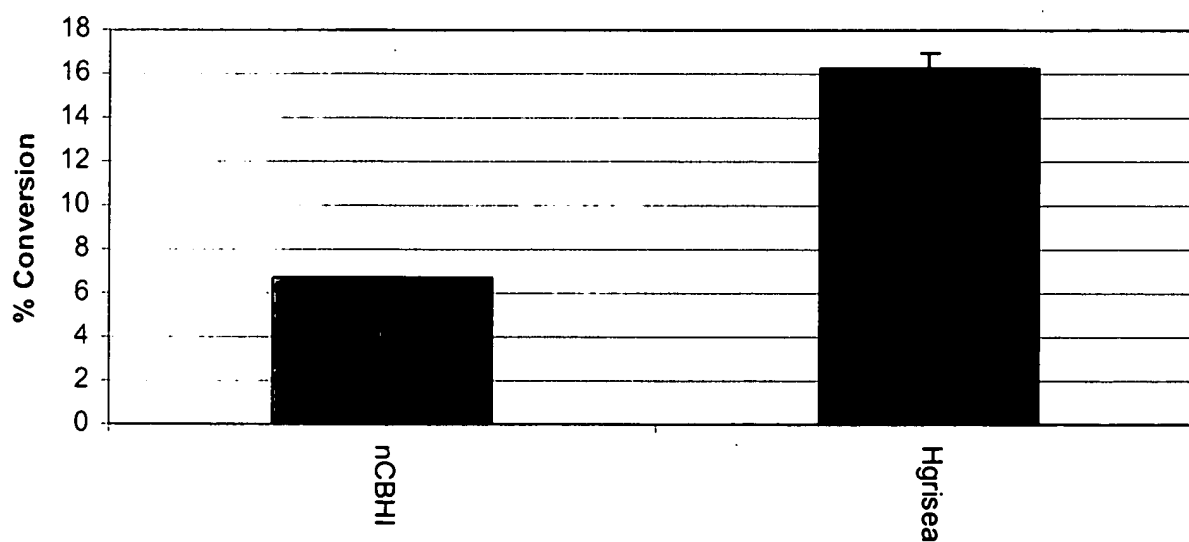


Figure 13A
3 fold rate difference b/n *H. grisea* and rCBHI

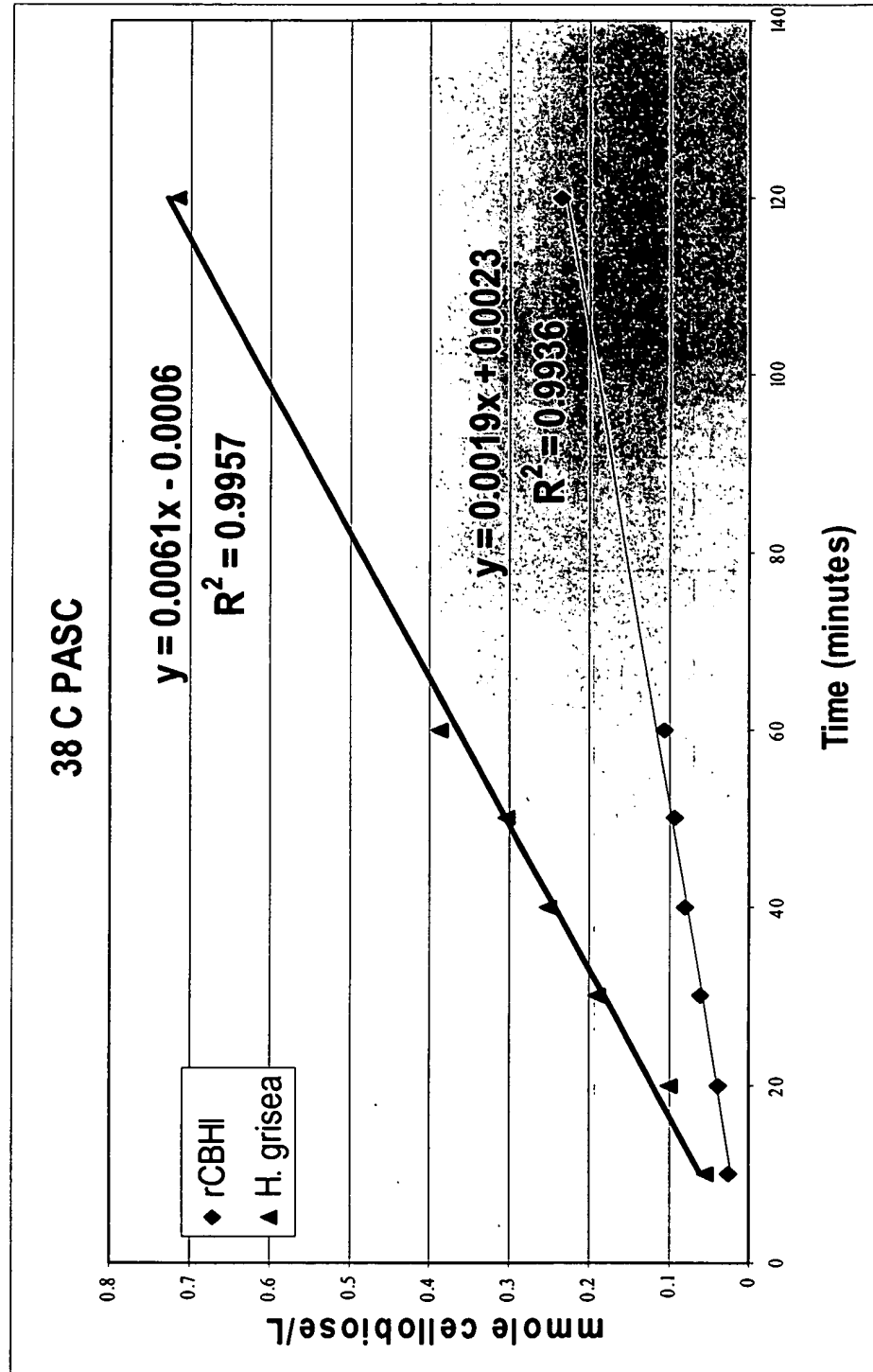


Figure 13B
4.8 fold rate difference

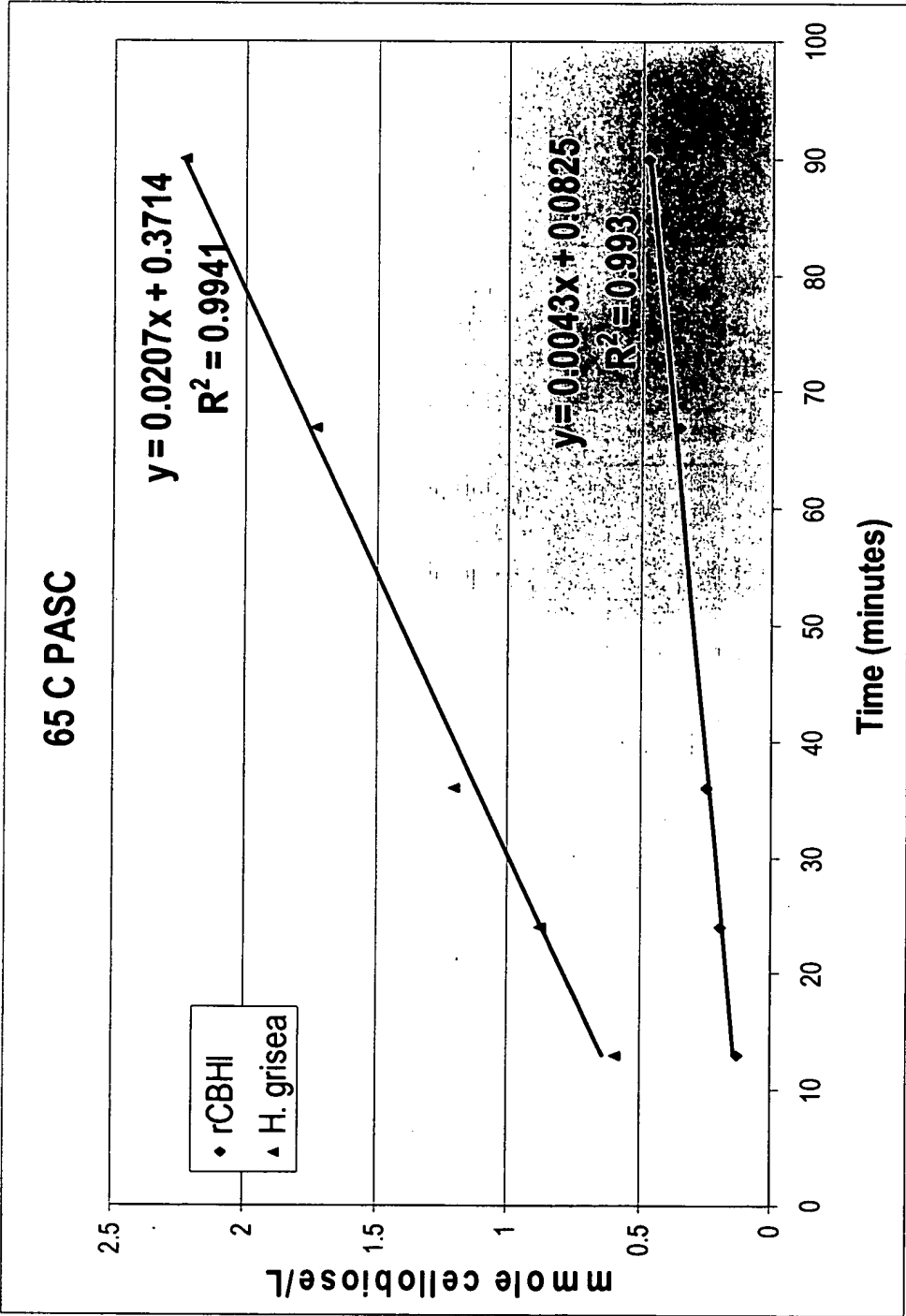
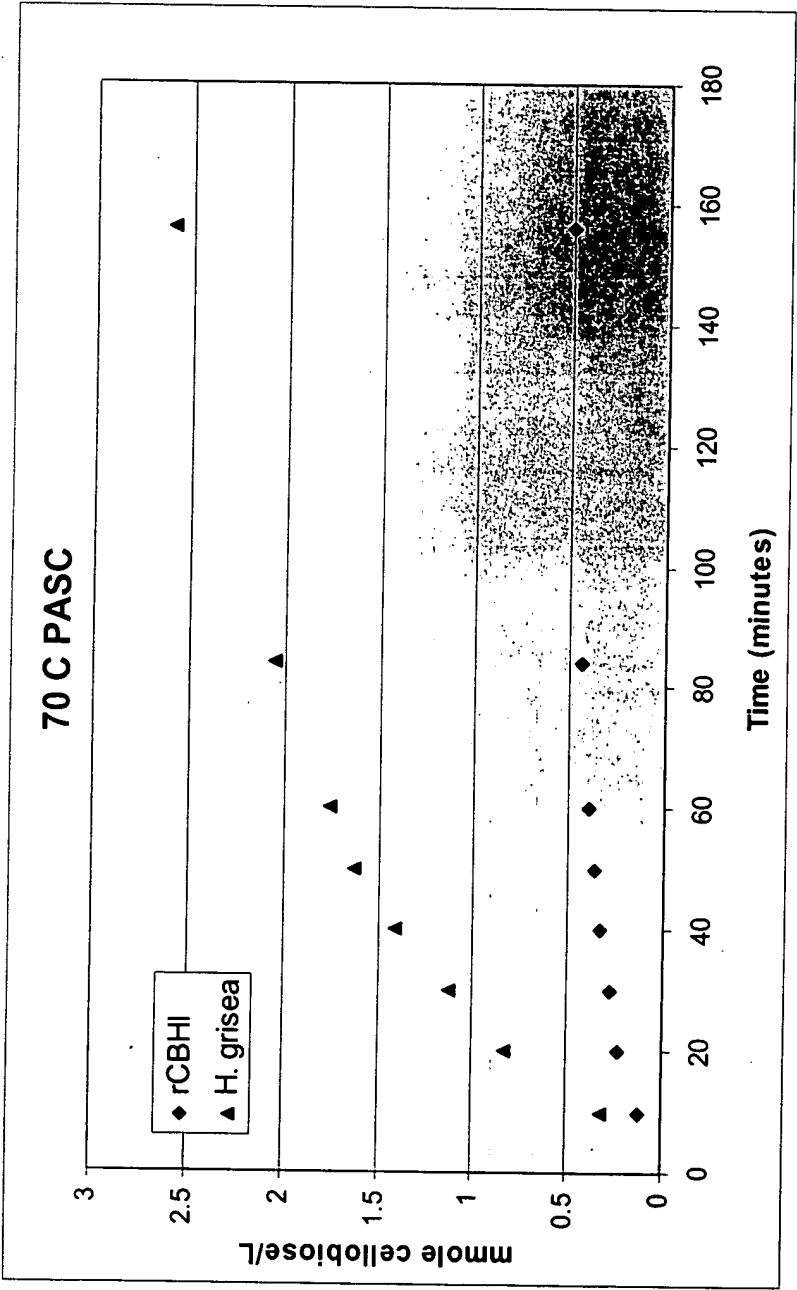


Figure 13C



```

1  ATGCGTACCG CCAAGTTCGC CACCCTCGCC GCCCTTGTGG CCTCGGCCGC
51 CGCCCAGCAG GCGTGCAGCC TCACCACCGA GAGGCACCCT TCCCTCTCCT
101 GGAAGAAGTG CACCGCCGGC GGCCAGTGCC AGACCGTCCA GGCTTCCATC
151 ACTCTCGACT CCAACTGGCG CTGGACTCAC CAGGTGTCTG GCTCCACCAA
201 CTGCTACACG GGCAACGAGT GGGATTCTAG CATCTGCACT GATGCCAAGT
251 CGTGCGCTCA GAACTGCTGC GTCGATGGTG CTGACTACAC CAGCACCTAT
301 GGCATCACCA CCAACGGTGA TTCCCTGAGC CTCAAGTTCG TCACCAAGGG
351 CCAGTACTCG ACCAACGTCG GCTCGCGTAC CTACCTGATG GACGGCGAGG
401 ACAAGTATCA GAGTAGGTTC TATCTTCAGC CTTCTCGCGC CTTGAATCCT
451 GGCTAACTTT TAACTTTCAC AGCCTTCGAG CTCCTCGGCA ACGAGTTCAC
501 CTTGATGTC GATGTCTCCA ACATCGGCTG CGGTCTCAAC GGCGCCCTGT
551 ACTTCGTCTC CATGGACGCC GATGGTGGTC TCAGCCGCTA TCCTGGCAAC
601 AAGGCTGGTG CCAAGTACGG TACCGGTAC TGCGATGCTC AGTGCCCCCG
651 TGACATCAAG TTCATCAACG GCGAGGCCAA CATTGAGGGC TGGACCGGCT
701 CCACCAACGA CCCCAACGCC GGCGCGGGCC GCTATGGTAC CTGCTGCTCT
751 GAGATGGATA TCTGGGAGGC CAACAACATG GCTACTGCCT TCACTCCTCA
801 CCCTTGCACT ATCATTGGCC AGAGCCGCTG CGAGGGCGAC TCGTGCAGGTG
851 GCACCTACAG CAACGACCGC TACGCCGGCG TCTGCGACCC CGATGGCTGC
901 GACTTCAACG CGTATCGCCA GGGCAACAAG ACCTTCTACG GCAAGGGCAT
951 GACCGTCGAC ACCACCAAGA AGCTACCCGT CGTACCCAG TTCCTCAAGG
1001 ACGCCAACGG CGATCTCGGC GAGATCAAGC GCTTCTACGT CCAGGATGGG
1051 AAGATCATCC CCAACTCCGA GTCCACCATC CCCGGCGTCG AGGGCAACTC
1101 CATCACCCAG GATTGGTGCG ACCGCCAGAA GGTTGCCTTT GGCGACATTG
1151 ACGACTTCAA CCGCAAGGGC GGCATGAAGC AGATGGGCAA GGCCCTCGCC
1201 GGCCCCATGG TCCTGGTCAT GTCCATCTGG GATGACCACG CCTCCAACAT
1251 GCTCTGGCTC GACTCGACCT TCCCTGTCTG TGCCGCTGGC AAGCCCGGCG
1301 CCGAGCGCGG TGCCTGCCCC ACCACCTCGG GTGTCCCTGC TGAGGTTGAG
1351 GCCGAGGCCC CCAACAGCAA CGTCGTCTTC TCCAACATCC GCTTCGGCCC
1401 CATCGGCTCG ACCGTTGCCG GCCTTCCCAG CGATGGCGGC AACAACGGCG
1451 GCAACACCAC CGTCCAGCCC CCGCCCAGCA CCACCACCAC CTCTGCCAGC
1501 AGCAGCACCA CCTCGGCTCC TGCCACCACC ACCACCGCCA GCGCTGGCCC
1551 CAAGGCTGGC CGCTGGCAGC AGTGCGGCGG CATCGGCTTC ACTGGCCCGA
1601 CCCAGTGCGA GGAGCCCTAC ACTTGACCA AGCTCAACGA CTGGTACTCT
1651 CAGTGCCTGT AA

```

Figure 14A
Scytalidium thermophilum CBH1 Genomic DNA

```

1  ATGCGTACCG CCAAGTTCGC CACCCTCGCC GCCCTTGTGG CCTCGGCCGC
   CGCCCAGCAG GCGTGCAGCC TCACCACCGA GAGGCACCCT TCCCTCTCCT
101 GGAAGAAGTG CACCGCCGGC GGCCAGTGCC AGACCGTCCA GGCTTCCATC
   ACTCTCGACT CCAACTGGCG CTGGACTCAC CAGGTGTCTG GCTCCACCAA
201 CTGCTACACG GGCAACGAGT GGGATTCTAG CATCTGCACT GATGCCAAGT
   CGTGCGCTCA GAACTGCTGC GTCGATGGTG CTGACTACAC CAGCACCTAT
301 GGCATCACCA CCAACGGTGA TTCCCTGAGC CTCAAGTTCG TCACCAAGGG
   CCAGTACTCG ACCAACGTCG GCTCGCGTAC CTACCTGATG GACGGCGAGG
401 ACAAGTATCA GACCTTCGAG CTCCTCGGCA ACGAGTTCAC CTTTCGATGTC
   GATGTCTCCA ACATCGGCTG CGGTCTCAAC GGCGCCCTGT ACTTCGTCTC
501 CATGGACGCC GATGGTGGTC TCAGCCGCTA TCCTGGCAAC AAGGCTGGTG
   CCAAGTACGG TACCGGCTAC TGCGATGCTC AGTGCCCCCG TGACATCAAG
601 TTCATCAACG GCGAGGCCAA CATTGAGGGC TGGACCGGCT CCACCAACGA
   CCCCAACGCC GGCGCGGGCC GCTATGGTAC CTGCTGCTCT GAGATGGATA
701 TCTGGGAGGC CAACAACATG GCTACTGCCT TCACTCCTCA CCCTTGCACT
   ATCATTGGCC AGAGCCGCTG CGAGGGCGAC TCGTGCGGTG GCACCTACAG
801 CAACGACCGC TACGCCGGCG TCTGCGACCC CGATGGCTGC GACTTCAACG
   CGTATCGCCA GGGCAACAAG ACCTTCTACG GCAAGGGCAT GACCGTCGAC
901 ACCACCAAGA AGCTCACCGT CGTACCCAG TTCCTCAAGG ACGCCAACGG
   CGATCTCGGC GAGATCAAGC GCTTCTACGT CCAGGATGGG AAGATCATCC
1001 CCAACTCCGA GTCCACCATC CCCGGCGTCG AGGGCAACTC CATCACCCAG
   GATTGGTGCG ACCGCCAGAA GGTTGCCTTT GGCGACATTG ACGACTTCAA
1101 CCGCAAGGGC GGCATGAAGC AGATGGGCAA GGCCCTCGCC GGCCCCATGG
   TCCTGGTCAT GTCCATCTGG GATGACCACG CCTCCAACAT GCTCTGGCTC
1201 GACTCGACCT TCCCTGTCGA TGCCGCTGGC AAGCCCGGCG CCGAGCGCGG
   TGCTTGCCCG ACCACCTCGG GTGTCCCTGC TGAGGTTGAG GCCGAGGCCG
1301 CCAACAGCAA CGTCGTCTTC TCCAACATCC GCTTCGGCCC CATCGGCTCG
   ACCGTTGCCG GCCTTCCCAG CGATGGCGGC AACAACGGCG GCAACACCAC
1401 CGTCCAGCCC CCGCCCAGCA CCACCACCAC CTCTGCCAGC AGCAGCACCA
   CCTCGGCTCC TGCCACCACC ACCACCGCCA GCGCTGGCCC CAAGGCTGGC
1501 CGCTGGCAGC AGTGCGGCGG CATCGGCTTC ACTGGCCCGA CCCAGTGC GA
   GGAGCCCTAC ACTTGACCA AGCTCAACGA CTGGTACTCT CAGTGCCTGT
1601 AA

```

Figure 14B
Scytalidium thermophilum CBH1 cDNA

1	MRTAKFATLAALVASAAAQQACSLTTERHPSLSWKKCTAGGQCQTVQASI	50
51	TLDSNWRWTHQVSGSTNCYTGNEWSSICTDAKSCAQNCCVDGADYTSTY	100
101	GITTNGDSLKSLKFVTKGQYSTNVGSRTYLMGDEKDYQTFELLGNEFTFDV	150
151	DVSNIGCGLNGALYFVSMDADGGLSRYPGNKAGAKYGTGYCDAQCPRIK	200
201	FINGEANIEGWTGSTNDPNAGAGRYGTCCSEMDIWEANNMATAFTPHPCT	250
251	IIGQSRCEGDSCGGTYSNDRYAGVCDPDGCDFNAYRQGNKTFYKGGMTVD	300
301	TTKKLTVVTQFLKDANGDLGEIKRFYVQDGKIIPNSESTIPGVEGNSITQ	350
351	DWCDRQKVAFGDIDDFNRKGGMKQMGKALAGPMVLVMSIWDDHASNMLWL	400
401	DSTFPVDAAGKPGAERGACPTTSGVPAEVEAEAPNSNVVFSNIRFGPIGS	450
451	TVAGLPSDGGNNGGNTTVQPPPSTTTTSASSSTTSAPATTTTASAGPKAG	500
501	RWQQCGGIGFTGPTQCEEPYTCTKLNDWYSQCL-	534

Figure 14C
Scytalidium thermophilum CBH1, including signal sequence

Figure 15: Alignment of the mature protein sequences for *Hypocrea jecorina* CBH1, *Humicola grisea* var. *thermoidea* CBH1.1, and *Scytalidium thermophilum* CBH

Genencor <i>Hypocrea jecorina</i> Cel7A	1	75
<i>Humicola grisea</i> CBH1.1	(1)	QSACTLQSETHPPLTWQCSSGGTCTQQTGSGVVIDANWRWTHATNSSTNCYDGN TS SLCPDNETCAKNCCLDG
<i>Scytalidium thermophilum</i> 69	(1)	QQA CS LT TER HP SL SNKCTAGGQCQTVAQASITLDSNWRWTHQVSGSTNCYTGNKWDTSICTDAKSCAQNCCVVG
Consensus	(1)	QQA CS LT TER HP SL SNKCTAGGQCQTVAQASITLDSNWRWTHQVSGSTNCYTGN EW DSICTDAKSCAQNCCVVG
Genencor <i>Hypocrea jecorina</i> Cel7A	76	150
<i>Humicola grisea</i> CBH1.1	(76)	AAYASTYGVTTSGNSLSIGFVTQSAQKN - VGARLYLMASD TT YQEF TL LGNEFSFDVDVSQ LP CG LN GALYFVSM
<i>Scytalidium thermophilum</i> 69	(76)	ADYTSTYGIT TNG DSLSLK FK VTGQ HS TN VS RTY LM DGEDKYQ TF ELLGNEFTFDVDSNIGCG LN GALYFVSM
Consensus	(76)	ADYTSTYGIT TNG DSLSLK FK VTGQ HS TN VS RTY LM DGEDKYQ TF ELLGNEFTFDVDSNIGCG LN GALYFVSM
Genencor <i>Hypocrea jecorina</i> Cel7A	151	225
<i>Humicola grisea</i> CBH1.1	(150)	DADGGVSKYPTNTAGAKYGTGYCDSCPRDLK FI NGQANVEGWEPSPNNANTGICGHGSCCSEMDIWEANSISEA
<i>Scytalidium thermophilum</i> 69	(151)	DADGGLSRYPGNKAGAKYGTGYCDAQCPRDIK FI NGEANI EG WTG ST NDPNAGAGRYGTC CS EMDIWEANNMATA
Consensus	(151)	DADGGLSRYPGNKAGAKYGTGYCDAQCPRDIK FI NGEANI EG WTG ST NDPNAGAGRYGTC CS EMDIWEANNMATA
Genencor <i>Hypocrea jecorina</i> Cel7A	226	300
<i>Humicola grisea</i> CBH1.1	(225)	LTPHPCTTVGQEI CE GDGCGGTYS DN RYGTC DP DCD WN PYRLGNTSFYGP GS FTLDTTK LT VVTQFETSG-
<i>Scytalidium thermophilum</i> 69	(226)	FTPH CT IIIGQSRCEGSCGGTYSNERYAGVCD PD GCDFN SR QGNKTFY GK - MTVD TT KKITVVTQFLK DAN
Consensus	(226)	FTPH CT IIIGQSRCEGSCGGTYSNERYAGVCD PD GCDFN SR QGNKTFY GK - MTVD TT KKITVVTQFLK DAN
Genencor <i>Hypocrea jecorina</i> Cel7A	301	375
<i>Humicola grisea</i> CBH1.1	(299)	----AINRYVYVQNGVT FQ QPN AE LSYSGNELND DY CTAEEAE FG SS - FSDK GG LTQ FK KATSGGMVLVMSLWD
<i>Scytalidium thermophilum</i> 69	(299)	GDLGEIKRFVYVQDGKII PN SESTIPGVEGNSITQ DW CD RQ KVAFGDID DF N RK GGMKMGKALAGP MV L VS IWD
Consensus	(301)	GDLGEIKRFVYVQDGKII PN SESTIPGVEGNSITQ DW CD RQ KVAFGDID DF N RK GGMKMGKALAGP MV L VS IWD
Genencor <i>Hypocrea jecorina</i> Cel7A	376	450
<i>Humicola grisea</i> CBH1.1	(369)	DYANMLWLDSTYPTNETS TP GAVR GS CS TS SGVPAQVESQSPNAKV TF SNIKFPIGSTGNPSG GNP -----
<i>Scytalidium thermophilum</i> 69	(374)	DHASNMLWLDSTFPVDAAG - KPGAERGACPTTSGVPAEVEAEAPNSNVWF SN IRFPIGSTVAGLPGAG--NGGN
Consensus	(376)	DHASNMLWLDSTFPVDAAG KPGAERGACPTTSGVPAEVEAEAPNSNVWF SN IRFPIGSTVAGLPGAG NGGN
Genencor <i>Hypocrea jecorina</i> Cel7A	451	518
<i>Humicola grisea</i> CBH1.1	(438)	PGGNPPG-----TTTTRRPATTGSSPGTQSHYQCGGICYS GP TV CA SGTTCQV LN PN Y YSQCL
<i>Scytalidium thermophilum</i> 69	(446)	NGGNPPP-----PTTTTSSAPATTTTASAGPKAGRWQCGGIGFTGPTQCEEPYICT KL NDWYSQCL
Consensus	(451)	PTTTTSSAPATTTTASAGPKAGRWQCGGIGFTGPTQCEEPYICT KL NDWYSQCL

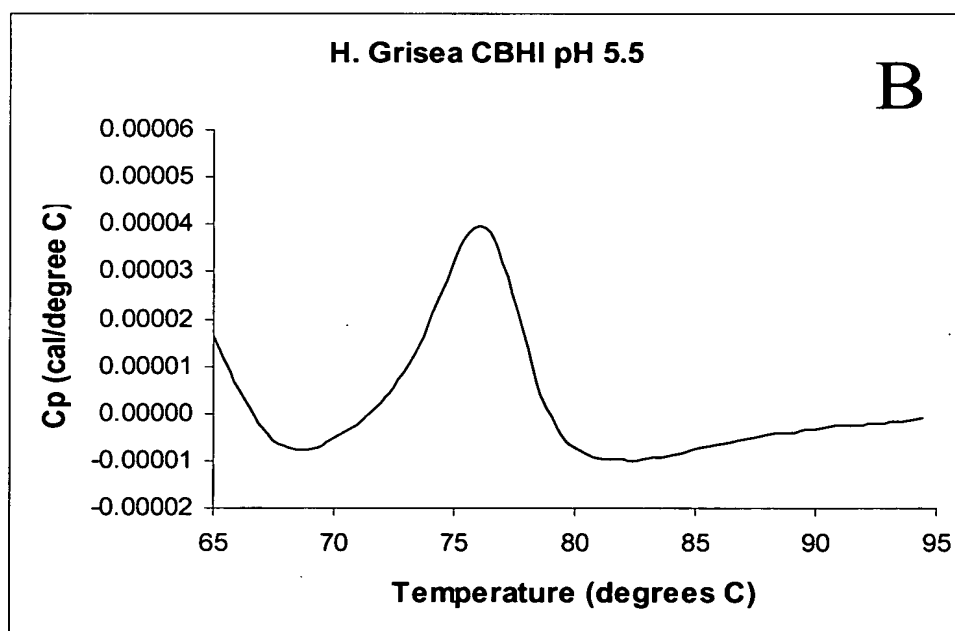
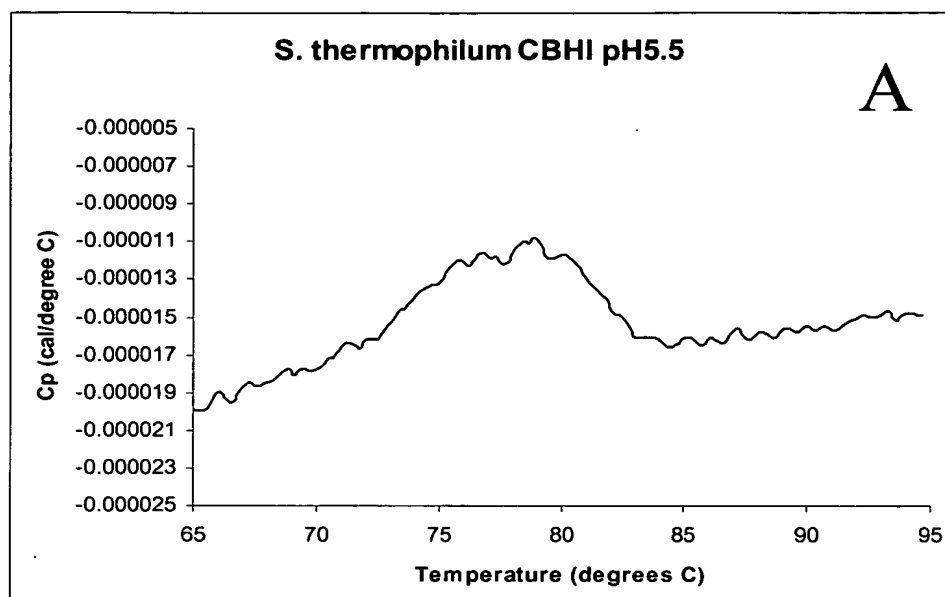


Figure 16